L'Actualité économique

La décomposition des mesures d'inégalité en sources de revenu : méthodes et applications

Stéphane Mussard

Volume 83, numéro 3, septembre 2007

URI : https://id.erudit.org/iderudit/018116ar
DOI : https://doi.org/10.7202/018116ar

Aller au sommaire du numéro

Éditeur(s)

HEC Montréal

ISSN

0001-771X (imprimé)
1710-3991 (numérique)

Citer cet article

Mussard, S. (2007). La décomposition des mesures d'inégalité en sources de revenu : méthodes et applications. *L'Actualité économique*, *83*(3), 415–445. https://doi.org/10.7202/018116ar

Résumé de l'article

La lecture de la littérature indique que la mesure des inégalités de revenus s'est largement développée depuis les années 1970. L'étude des inégalités mesurées sur les revenus des individus est nécessaire mais non suffisante pour appréhender la complexité des déterminants des inégalités. En ce sens, les techniques de décompositions des mesures d'inégalité en sources de revenu sont intéressantes. Elles permettent de mettre en évidence de nouveaux indices statistiques dont la structure autorise l'analyse des sources de rémunération (salaires, primes, taxes, pensions, *etc.*), les corrélations de ces sources aux rangs des individus dans la société ou leurs parts dans le revenu moyen. Notre analyse s'effectue autour de la décomposition de la mesure de Gini, les débats qu'elle a pu susciter et les améliorations de la méthode en partant de la notion de pseudo-Gini à celle de Gini étendu. Les dernières méthodes en date sont aussi exposées afin de mettre en exergue les possibilités de généralisation en ce domaine en utilisant soit l'analyse économétrique soit la valeur de Shapley.

L'Actualité économique, Revue d'analyse économique, vol. 83, n° 3, septembre 2007

*La décomposition des mesures d'inégalité en sources de revenu : méthodes et applications**

Stéphane MUSSARD
LAMETA, Université Montpellier I,
GRÉDI,
GEREM

RÉSUMÉ – La lecture de la littérature indique que la mesure des inégalités de revenus s'est largement développée depuis les années 1970. L'étude des inégalités mesurées sur les revenus des individus est nécessaire mais non suffisante pour appréhender la complexité des déterminants des inégalités. En ce sens, les techniques de décompositions des mesures d'inégalité en sources de revenu sont intéressantes. Elles permettent de mettre en évidence de nouveaux indices statistiques dont la structure autorise l'analyse des sources de rémunération (salaires, primes, taxes, pensions, *etc.*), les corrélations de ces sources aux rangs des individus dans la société ou leurs parts dans le revenu moyen. Notre analyse s'effectue autour de la décomposition de la mesure de Gini, les débats qu'elle a pu susciter et les améliorations de la méthode en partant de la notion de pseudo-Gini à celle de Gini étendu. Les dernières méthodes en date sont aussi exposées afin de mettre en exergue les possibilités de généralisation en ce domaine en utilisant soit l'analyse économétrique soit la valeur de Shapley.

ABSTRACT – The literature indicates that income inequality measures have been developed since the 70's. In order to apprehend the underlying components of the inequality measures, it is necessary but not sufficient to study overall incomes. In this respect, decomposing inequality indices by income source is welcome. The methodology brings out new statistical indices, for which it is possible to analyze factor components (wages, fringe benefits, taxes, pensions, *etc.*), correlations between the income sources and the ranks of individuals within the society, and the contribution of each source in the mean income. The paper draws an analysis of the Gini decomposition by income source, the debates between researchers, and the development of the decomposition technique, starting from the concept of pseudo-Gini to that of extended Gini. The last approaches of the literature are exposed in order to point out the possibilities of generalization, using either econometric models or the Shapley value.

* L'auteur tient à remercier le ministère de le Recherche du Luxembourg, le CEPS/INSTEAD, Philippe Van Kerm et deux arbitres anonymes pour la qualité de leurs commentaires.

Introduction

Les inégalités dans la répartition des revenus et celle des richesses doivent-elles être réduites? Plusieurs économistes l'ont soutenu. Pareto et Keynes par exemple. Pareto (1896) préconise une modification de la distribution des revenus au détriment des plus riches et en faveur des plus pauvres. Dans le dernier chapitre de la Théorie générale, Keynes (1936) n'hésite pas à conclure que sa conception du fonctionnement des économies de marché justifie une réduction sensible des inégalités arbitraires et inéquitables qui caractérisent la distribution des revenus et la répartition des richesses au sein du monde dans lequel il vit. Ces prises de position permettent de comprendre que la réflexion sur la mesure des inégalités a pu constituer et constitue l'un des thèmes de recherche privilégié des économistes.

En ce domaine, les travaux de Lorenz (1905) et ceux de Gini (1912, 1914, 1916) peuvent être considérés comme précurseurs. Ils constituent les points de départ des recherches développées ultérieurement par Kolm (1966), Atkinson (1970) et Sen (1973). Toutefois, ce sont les résultats obtenus en 1969 par Rao qui vont retenir notre attention. Dès 1969, Rao essaie de comprendre les rouages des mécanismes de décomposition des mesures d'inégalité en privilégiant soit une décomposition en sous-groupes soit une décomposition en sources de revenu. La première est consacrée à la détermination des inégalités à l'intérieur des groupes et entre les groupes. Cette technique permet notamment de mesurer les distances qui séparent certains types de ménage (*Cf.* Dagum, 1997). La seconde, celle qui fait l'objet de notre réflexion, appelée décomposition en sources de revenu ou encore décomposition en facteurs, s'intéresse à la contribution des différentes sources de revenu (revenus du travail, revenu du capital, prestations sociales, taxes, *etc.*) à la mesure de l'inégalité globale. Se démarquant de Rao (1969), et des nombreuses applications que sa méthode a permises, Shorrocks (1982) va proposer un nouvel outil. Les résultats auxquels il parvient en utilisant la démarche axiomatique conduisent à douter de la crédibilité de la décomposition de la mesure de Gini en facteurs, alors même que le nombre de publications privilégiant cette technique constitue un indice authentique de l'engouement qu'elle a suscité. Parmi ces publications, citons celles qui font le lien avec la croissance et le développement économique (Fei, Ranis et Kuo, 1978), le lien avec la consommation (Garner, 1993) ou encore l'analyse économétrique (Murdoch et Sicular, 2002) et celles dont l'apport nous paraît le plus important, Chantreuil et Trannoy (1999) et Shorrocks (1999). Ces recherches donnent l'occasion d'employer la valeur de Shapley, concept central de la théorie des jeux coopératifs, pour généraliser la décomposition en sources de revenu.

Dans les lignes qui suivent, nous ne démontrons pas le lien qui existe entre inégalités de revenu et inéquités, mais nous admettons que la réduction des inégalités de revenu constitue un moyen permettant de ralentir, bien que ceci ne soit pas systématique, des processus inéquitables tels que, par exemple, l'inégalité des chances. En effet certains types d'inégalités de revenu, comme les inégalités salariales, peuvent être à l'origine de mécanismes incitatifs favorisant l'effort au travail et ainsi la croissance économique. Inversement, la baisse des inéquités engendre la baisse des inégalités de revenu, mais l'implication n'est guère, là encore, systématique. En revanche, il est nécessaire de lutter contre les inégalités inéquitables. Les formes d'interventions introduisent des distorsions dans la détermination de la rémunération des facteurs de production, mais elles sont nécessaires à l'égalisation des chances, favorisant par exemple l'insertion des femmes et des jeunes sur le marché du travail (*Cf.* WDR, 2006).

Afin d'apprécier l'ampleur de l'évolution des techniques de décomposition en facteurs, nous privilégions une approche chronologique. Nous montrons tout d'abord que la première décomposition statistique de l'indice de Gini, réalisée par Rao (1969) (première section), fut le point de départ de nombreux développements (deuxième section), avant d'être remise en cause (troisième section). Nous rappelons (quatrième section) que Lerman et Yitzhaki (1985) ont conçu une nouvelle approche, qu'ils qualifient de désirable, dans laquelle la mesure de Gini décomposée permet de rendre compte de l'aversion des individus à l'égard des inégalités. Enfin, avant de conclure sur l'état actuel de nos connaissances en matière de décomposition en sources de revenu, en soulignant notamment (sixième section) que le domaine n'a pas atteint sa maturité, nous présentons la généralisation de la décomposition en facteurs à toutes les mesures d'inégalité (cinquième section).

1. La genèse de la décomposition en facteurs : Rao (1969)

La paternité de la décomposition de la mesure de Gini en sources de revenu peut être attribuée à Rao (1969). Après avoir spécifié une décomposition matricielle de l'indice de concentration en sous-groupes, l'auteur exprime dans la deuxième partie de sa réflexion, une méthode permettant de décomposer l'indice de concentration en contributions représentant le poids des différentes sources de rémunération. Rappelons, qu'une courbe de concentration permet le calcul de l'indice de concentration. Ce dernier autorise la mesure des inégalités de revenus, de consommations ou plus généralement des inégalités issues d'une variable x lorsque les individus sont classés par ordre d'importance de revenus. La courbe de Lorenz permet quant à elle de mesurer les inégalités issues de x (*via* la mesure de Gini) lorsque les individus sont classés par ordre d'importance de la variable x[1].

Soit une population composée de n individus, $i \in \{1, ..., n\}$. Les revenus des individus comprennent q sources de revenu indexées par $m \in \{1, ..., q\}$. Soit la fonction $R_{(mi)}$. Elle exprime le rang du i^e individu détenant la source de revenu m, lorsque les sources de nature m (pour tout m) sont ordonnées de manière croissante. Parallèlement, on peut définir le rang de l'individu i lorsque les individus sont ordonnés de manière croissante selon leurs revenus : $R_{(ti)}$. Avant 1980, les

1. Les différences et les liens unissant l'indice de concentration et l'indice de Gini sont étudiés par Pyatt, Chen et Fei (1980), *Cf.* section 2.3 *infra*). Par exemple, l'indice de concentration peut être négatif, notamment lorsque les classes sont ordonnées de manière décroissante.

inégalités sont étudiées en considérant des classes de revenu. Nous verrons par la suite la conséquence fondamentale de ce découpage particulier (*Cf.* section 2.3 *infra*). On considère que les distributions de chaque source de revenu m sont étudiées selon k classes S_{jm} définies par l'ordre croissant $R_{(mi)}$ (on a donc k quantiles : $j \in \{1, ..., k\}$ comprenant $\frac{n}{k}$ individus). La distribution du revenu total est aussi divisée en k classes S_{jt} comprenant chacune $\frac{n}{k}$ individus et où les individus sont classés selon le critère $R_{(ti)}$. On note $\tilde{Q}$ (Q_{jm}) la somme cumulée des proportions (de revenu total) de sources m inhérentes à chaque classe S_{jm} (S_{jt}). Soit la somme cumulée Q_{jt} (sur les classes S_{jt}, non S_{jm}) des proportions de revenu total inhérentes à chaque classe S_{jt}. Par conséquent, $\sum_m Q_{jm} = Q_{jt}$. Soit le terme $Q_{.m}$ désignant la proportion de revenu total imputable au facteur m et P_{jt} la somme cumulée des proportions d'individus selon les classes S_{jt}. L'indice de concentration s'exprime par :

$$c = \sum_{j=1}^{k-1} (P_{jt} Q_{(j+1)t} - P_{(j+1)t} Q_{jt}) . \tag{1}$$

De la même manière, l'indice de concentration associé à la source m est :

$$c_m = \sum_{j=1}^{k-1} (P_{jm} \tilde{Q}_{(j+1)m} - P_{(j+1)m} \tilde{Q}_{jm}) . \tag{2}$$

Soit le ratio $\frac{Q_{jm}}{Q_{.m}} =: Q^*_{jm}$. En appliquant la formule (1) aux variables P_{jt} et Q^*_{jm}, on obtient un indice de concentration modifié :

$$c^*_m = \sum_{j=1}^{k-1} (P_{jt} Q^*_{(j+1)m} - P_{(j+1)t} Q^*_{jm}) . \tag{3}$$

Par rapport à la mesure c_m, le coefficient c^*_m est atypique car son calcul s'effectue en tenant compte des proportions d'individus (P_{jt}) en lieu et place des proportions (P_{jm}). De plus, c^*_m est obtenu à partir du classement S_{jt} alors que c_m est obtenu à partir des rangs S_{jm}.

La décomposition de l'indice de concentration en sources de revenu s'exprime alors sous la forme suivante :

$$c = \mathbb{Q}'_{.m} \mathbb{C}^*_m \tag{4}$$

où $\mathbb{Q}'_{.m}$ est le vecteur colonne transposé des proportions des sources m dans le revenu global et où $\mathbb{C}^*_m$ est le vecteur colonne des q indices de concentration modifiés (c^*_m) inhérents aux q sources de revenu. Or, étant donné que la décomposition repose sur la mesure c^*_m, Rao reformule l'expression précédente afin de proposer une décomposition exacte et conforme à l'indice c_m :

$$c = \mathbb{Q}'_m \mathbb{C}_m - \mathbb{Q}'_m \mathbb{C}_m f_m \tag{5}$$

où $\mathbb{C}_m$ est le vecteur colonne des indices de concentration des sources m, où $f_m = 1 - \frac{c_m^*}{c_m}$, $f_m \in [0, 2]$ et où $\mathbb{C}_m f_m$ est le vecteur colonne des produits scalaires $c_m \cdot f_m$. Lorsque les classes S_{jt} et S_{jm} son identiques $\forall m \in \{1, ..., q\}$: $c_m^* = c_m$ alors $f_m = 0 \Longrightarrow c = \mathbb{Q}'_m \mathbb{C}_m$. L'indice de concentration global devient donc une moyenne des indices de concentration, de chaque facteur, pondérés par le poids de ces facteurs. Une autre condition plus forte permet d'atteindre ce résultat : $R_{(mi)} = R_{(ti)}$, $\forall m \in \{1, ..., q\}$. Autrement dit, les individus doivent avoir le même rang selon la distribution de revenu globale et selon la distribution de chacun des facteurs. Il s'agit d'une condition nécessaire mais non suffisante pour respecter la condition $S_{jt} = S_{jm}$ et obtenir le résultat $c = \mathbb{Q}'_m \mathbb{C}_m$. Empiriquement, ce résultat est plus difficile à atteindre que la relation (5) qui indique que si l'on accède à la parfaite répartition au sein des distributions de chacune des q sources de revenu, alors l'inégalité globale est nulle. Le premier terme $\mathbb{Q}'_m \mathbb{C}_m$ est considéré comme la valeur maximale que c peut atteindre. Le deuxième $\mathbb{Q}'_m \mathbb{C}_m f_m$ mesure l'intensité avec laquelle les inégalités issues de chaque source peuvent s'annuler. En effet, $c = 0$ lorsque $c_m = 0$ ou lorsque le second terme annule l'effet du premier. Les différentes interprétations peuvent être classées par rapport aux valeurs prises par f_m :

- Si $f_m = 0$, le deuxième élément $\mathbb{Q}'_m \mathbb{C}_m f_m$, mesurant les inégalités provenant des sources autres que m, ne permet pas d'atténuer les inégalités induites par la source m. Il s'agit d'une inégalité de « supplément ».
- Si $f_m = 2$, les inégalités impulsées par les sources (excepté m) viennent contrer l'effet des inégalités issues du facteur m.
- Si $f_m = 1$, les deux éléments se neutralisent.

Morduch et Sicular (2002) vont critiquer cette technique (*Cf.* section 5.1 *infra*) car elle ne permet pas de capter l'effet d'une source constante sur l'inégalité totale (*Cf.* aussi Araar, 2006). Mais l'approche de Rao (1969) reste fondamentale. Elle repose en premier lieu sur une approche matricielle. Silber (1989) estime, à ce propos, que les approches décomposées sont souvent basées sur des algorithmes intéressants mais encore difficiles à estimer, même en utilisant l'outil informatique. Il propose, comme l'a fait Rao, de simplifier la décomposition en sources de revenu en proposant une approche matricielle (*Cf.* aussi Silber, 1993). Deuxièmement, la méthode de Rao procure deux types de décomposition : l'une est basée sur l'indice de concentration c_m inhérent à chaque source de rémunération, l'autre est un indice de concentration « composite », que la littérature nommera pseudo-Gini.

2. Les applications de l'approche de Rao (1969)

L'approche développée par Rao (1969) a constitué la pierre angulaire de nombreuses méthodes de décomposition, notamment la décomposition de l'indice de

Gini en sources de revenu. Dans les lignes qui suivent, l'indice de Gini est tout d'abord privilégié (sections 2.1 et 2.2), puis l'indice de Gini est construit comme étant un cas particulier de l'indice de concentration (section 2.3).

2.1 *Le pseudo-Gini de Fei, Ranis et Kuo (1978) : le lien avec l'économie du développement*

Fei, Ranis et Kuo (que nous noterons désormais FRK) utilisent la décomposition de la mesure de Gini à des fins spécifiques : celles de l'étude de la croissance économique, du développement économique et de la détermination des sources d'inégalité qui influencent le comportement des familles riches et celui des familles pauvres. La démarche des auteurs est empirique, elle se base sur le développement purement théorique de Rao (1969).

FRK étendent l'idée de Kuznets (1955) selon laquelle un lien étroit unit les inégalités de revenu et la croissance économique[2]. Comment la croissance économique s'est-elle aussi rapidement accrue? Pour répondre à cette question, l'analyse de l'histoire économique d'un pays peut être envisagée. Une alternative consiste à appréhender les déterminants de la distribution des revenus. On considère pour cela une distribution constituée de n ménages $(x_1, ..., x_i , ..., x_n)$, où chaque ménage i possède un nombre fini de sources de revenu telles que : $x_i = x_i^1 + x_i^2 + ... + x_i^m + ... + x_i^q$. On note G, l'indice de Gini mesurant les inégalités de revenu globales et G_m l'indice de Gini inhérent à la source de revenu m. De manière intuitive, l'indicateur global peut se décomposer à l'aide des indices G_m. Il suffit d'envisager une moyenne pondérée telle que : $G \cong \phi_m G_m$, où ϕ_m est la part moyenne de la source m dans le revenu total $\left(\dfrac{\mu^m}{\mu}\right)$ et où μ^m est la moyenne arithmétique de la source de m et μ celle du revenu global. Ainsi, en s'appuyant sur une approximation, une décomposition cohérente peut s'exprimer par :

$$G = \hat{G} - \theta \tag{6}$$

où

$$\hat{G} = \sum_{m=1}^{q_1} \phi_m G_m + \sum_{m=1}^{q_2} \phi_m G_m + \sum_{m=1}^{q_3} \phi_m G_m, \forall \sum_{m} \phi_m = 1 \tag{7}$$

et où θ représente un terme d'erreur. Les facteurs de revenu sont regroupés en trois types. Les sources « type 1 » (indexées par m et au nombre de q_1) sont les facteurs détenus en plus grande quantité (de manière absolue et relative) par les

2. Kuznets construit une courbe en U renversée montrant que les inégalités de revenu augmentent lorsqu'un pays se développe (augmentation du revenu par tête), puis les inégalités diminuent. L'augmentation s'explique notamment par le fait que la croissance profite dans un premier temps à une faible proportion de la population.

ménages riches. Ces revenus sont liés à la propriété privée (les biens matériels, fonciers, *etc*.). Les facteurs « type 2 » (au nombre de q_2) représentent les sources qui sont proportionnellement moins importantes chez les familles pauvres comme le salaire. Enfin, les facteurs « type 3 » (au nombre de q_3 et associés au signe –) décroissent de manière absolue pour les familles qui sont les plus dotées en bien-être. Ces sources sont des transferts de revenu. La classification de ces trois types de source de revenu permet d'élaborer un modèle estimé sur un panel de Taiwan (1964, 1966, 1968, 1970, 1971 et 1972). Le principal résultat de l'estimation économétrique de l'équation (7) est la faiblesse des sources de type 3.

Ce modèle est cependant beaucoup plus apte à rendre compte du lien qui prévaut entre les inégalités et la croissance économique. Considérons une économie composée d'une seule branche et de deux facteurs de richesse ($q = 2$) : le capital (K) et le travail (L). Par conséquent, l'indice de Gini global peut se réécrire comme suit :

$$G = \phi_k G_k + \phi_l G_l, \forall \sum_m \phi_m = 1, \theta \to 0 \tag{8}$$

où ϕ_k et ϕ_l sont respectivement les parts de capital et de travail. La différentielle de l'expression précédente permet de déterminer deux impacts marginaux :

$$G = \underbrace{(G_l - G_k)\frac{d\phi_l}{dt}}_{D} + \underbrace{\phi_l \frac{G_l}{dt} + \phi_k \frac{G_k}{dt}}_{I}. \tag{9}$$

La statique comparative permet d'expliquer le changement du Gini total (G) en fonction des effets D et I. Le premier effet (D) est un effet de distribution fonctionnelle. Il décrit le changement de l'inégalité totale par rapport aux variations des parts de chaque facteur : capital et travail. Le second (I) est l'effet des facteurs de Gini. Il représente les changements de G au cours du temps par l'intermédiaire des impacts favorables ou défavorables des indices de Gini pour chaque facteur. Lorsque les salaires sont répartis de manière plus égalitaire que les capitaux ($G_l < G_k$), et lorsque les salaires sont favorisés (ϕ_l augmente), alors l'effet de distribution fonctionnelle représente une amélioration de la distribution de revenus des familles. Afin d'évaluer l'impact de chaque facteur, FRK introduisent le concept de pseudo-Gini en utilisant la courbe de Lorenz[3]. Le coefficient de Gini est égal à $G = 1 - 2A$, où A est l'aire située entre la courbe de Lorenz et l'axe des abscisses[4]. L'aire peut s'exprimer de la manière suivante :

3. La courbe de Lorenz s'obtient en liant les points de coordonnée (P_i, Q_i) dans le plan (P, Q) où P_i désigne la proportion de population ayant un rang (selon le revenu) inférieur à i et où Q_i représente la part de revenu détenue par P_i dans le revenu global.

4. Pour mesurer cette aire, il suffit d'utiliser la formule de l'aire d'un trapèze : $\dfrac{(q_i + q_{i-1})\frac{1}{n}}{2}$. En prenant la somme des ces aires selon i, on trouve la relation (10).

$$A = 1 - \frac{\sum_{i=1}^{n} iq_i}{n} + \frac{1}{2n} \tag{10}$$

où q_i représente la proportion (non cumulée) de revenu total détenue par la proportion de population (ou l'individu) situé au rang i : $q_i = \frac{x_i}{n\mu}$, telle que $\sum_{i=1}^{n} q_i = 1$. Cette expression repose sur une transformation monotone non décroissante des poids (i) et des proportions de revenus (q_i). L'indice se réécrit :

$$G = \frac{2}{n} u_x - \frac{n+1}{n} \tag{11}$$

où $u_x = \sum_{i=1}^{n} iq_i$. Ensuite, définissons q_i^m comme étant la part de la source m totale détenue par l'individu i. Si le terme $\bar{u}_m$ est un « pseudo-indicateur » de u_x lorsqu'on s'intéresse à la source m, autrement dit $\bar{u}_m = \sum_{i=1}^{n} iq_i^m$, alors l'indice de Gini peut s'exprimer par :

$$G = \sum_{m=1}^{q} \left[\frac{2}{n} \bar{u}_m - \frac{n+1}{n} \right] = \sum_{m=1}^{q} \phi_m \bar{G}_m . \tag{12}$$

Les pseudo-Gini $\bar{G}_m$ indiquent que la condition de monotonie concernant les poids est perdue. En effet, les éléments q_i^m du terme $\bar{u}_m$ ne sont pas classés par ordre croissant. Ils sont ordonnés selon le classement croissant des revenus : $x_1 \le x_2 \le ... \le x_n$. Dans ces conditions, les éléments q_i^m ne respectent pas systématiquement cette même condition de monotonie. On retrouve alors la discussion de Rao (1969) sur la différence qui prévaut entre les rangs associés au vecteur de revenu global et ceux associés au vecteur de chaque source de revenu, et les conditions qu'il faut réunir afin que le pseudo-Gini de la source m ($\bar{G}_m$) soit égal au véritable indice de Gini mesuré sur la source m (G_m).

En définitive, cette approche est équivalente à celle de Rao (1969), mais néanmoins plus complète, puisque FRK mettent en exergue une décomposition à des fins économiques comme la détermination des sources d'inégalité expliquant les disparités qui règnent entre les pauvres et les non-pauvres.

2.2 *Une application aux ménages colombiens : Fields (1979)*

La méthodologie adoptée par Fields provient de celle de FRK (1978). L'indice de Gini global calculé sur un ensemble de sources de rémunération s'exprime sous la forme d'une moyenne pondérée des pseudo-Gini. L'auteur étudie les sources de revenu de 2 949 ménages colombiens de 1967 et 1968. Le revenu total de chaque ménage est constitué de quatre facteurs principaux. La première source concerne les revenus salariaux. Il s'agit des salaires, du partage des bénéfices et des primes. Le deuxième facteur concerne les revenus indépendants issus de services rendus ou de travaux professionnels. Le troisième facteur concerne les revenus du capital : intérêts, dividendes et rentes. La dernière source représente les transferts de revenu. En rappelant que $\phi_m = \frac{\mu^m}{m}$ où μ^m est la moyenne des source de type m, l'indice de Gini global s'exprime alors sous la forme d'une moyenne pondérée de quatre pseudo-Gini :

$$G = \phi_1 \bar{G}_1 + \phi_2 \bar{G}_2 + \phi_3 \bar{G}_3 + \phi_4 \bar{G}_4. \tag{13}$$

D'après la décomposition de FRK (1978), le pseudo-Gini d'une source m ($\bar{G}_m$) est une fonction de l'indice de Gini associé à la source m (G_m) :

$$\bar{G}_m = G_m \, \Omega_m \tag{14}$$

où Ω_m est un coefficient de corrélation relatif proche de celui de Pearson. Cependant, son dénominateur est différent. Il est en effet basé sur la covariance entre le facteur m et le rang des individus selon m (non sur le produit des écarts-types). Soit une population de n ménages (i = 1, ..., n), avec x^m un vecteur de n éléments représentant la m^e source de revenu. On note R le vecteur des positions (rangs) de chaque ménage en fonction du classement croissant de leur revenu global et R_m le vecteur de leur rang en fonction du classement croissant des éléments de x^m. Par conséquent, le coefficient Ω_m est défini par :

$$\Omega_m = \frac{\mathrm{Cov}(x^m, R)}{\mathrm{Cov}(x^m, R^m)} \tag{15}$$

où $\mathrm{Cov}(x^m, R)$ est la covariance entre la source m et la position relative de chaque ménage dans la distribution de revenus globale. D'autre part, $\mathrm{Cov}(x^m, R^m)$ représente la covariance entre la source m et le rang des ménages lorsqu'ils sont classés selon la source x^m. Par conséquent, la décomposition de l'indice de Gini devient :

$$G = \sum_{m=1}^{q} \phi_m \Omega_m G_m \,. \tag{16}$$

Outre cette brève analyse des corrélations (qui sera reprise et développée par Lerman et Yitzhaki (1985), *Cf.* section 4 *infra*), il est possible de s'intéresser à la contribution relative de chaque source de revenu dans la mesure du Gini total :

$$100\,\% = \sum_{m=1}^{q} \frac{\phi_m \Omega_m G_m}{G} \,. \tag{17}$$

La réécriture de la décomposition de FRK permet ainsi de mesurer, dans le cadre de la réflexion de Fields (1979), l'influence qu'exercent les quatre sources de revenu colombiennes. Les facteurs issus du travail, salaires et revenus indépendants (sources 1 et 2), représentent respectivement 27,02 % et 42,08 % de l'inégalité totale. Les revenus provenant du capital génèrent 26,47 % de l'inégalité contre 4,42 % pour les transferts de revenus.

L'auteur soulève ensuite une question capitale. Existe-t-il une différence significative entre les estimations effectuées à partir de données individuelles et celles menées à partir de données groupées? Les revenus sont répartis en 21 classes. Le tableau 1 expose, pour chaque méthode d'estimation, les contributions de chaque source à l'inégalité totale.

TABLEAU 1

CONTRIBUTIONS DES SOURCES À L'INÉGALITÉ TOTALE

Contributions (%)	**Données groupées**	**Données individuelles**
S^1	27,65	27,02
S^2	41,87	42,08
S^3	25,46	26,47
S^4	4,48	4,42

SOURCE : PRESFAM, Center for the Study of Economic Development, University of the Andes, Bogota

Les résultats précédemment obtenus par FRK sur les revenus de Taiwan et du Pakistan viennent ainsi confirmer les résultats obtenus par Fields (1979). En effet, les revenus du travail sont répartis de manière très inégalitaire. Ils accroissent les inégalités globales. Cette conclusion, autant valable pour les données individuelles que pour les données groupées, est en partie due à l'importance du poids des sources 1 et 2 (salaires et revenus indépendants) dans le revenu total.

En somme, l'auteur constate simplement que les écarts entre les contributions sont très faibles, alors que cette conclusion n'est plus valable en utilisant les indicateurs de Gini par source de revenu (G_m) et les corrélations (Ω_m). Pyatt, Chen, et Fei (1980) vont pourtant démontrer qu'une différence fondamentale prévaut entre les indices (non les contributions) lorsqu'ils sont estimés à partir de données individuelles en lieu et place de données groupées.

2.3 *Estimations sur données individuelles et données groupées : Pyatt, Chen et Fei (1980)*

Avant de mettre en évidence les difficultés auxquelles les économistes se heurtent en interprétant des mesures effectuées sur données individuelles ou bien sur données groupées, les auteurs proposent de démontrer que la mesure de Gini est un cas particulier du coefficient de concentration.

Reprenons la même fonction qu'utilise Rao (1969) pour exprimer le rang des ménages (individus). La règle de classification est t. Lorsque t_i est le plus petit élément d'un ensemble de n éléments ($i \in \{1, 2, ..., n\}$) : $R_{(ti)} = 1$ et lorsque t_i est le plus grand : $R_{(ti)} = n$, n étant le nombre de ménages. Un problème qui n'avait jamais été évoqué jusqu'ici est soulevé. Il peut en effet exister plusieurs ménages possédant le même rang dans la distribution. Il faut donc leur attribuer un rang identique. On leur octroie la moyenne arithmétique de leur rang. De cette manière, il est possible de définir le rang moyen de la distribution :

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_{(t_i)} = \frac{n+1}{2}. \tag{18}$$

On considère que x_i est une variable inconnue, mais on exige que la moyenne de cette variable μ soit non négative, bien que les x_i puissent l'être. Rappelons que q_i est la proportion de revenu total (non cumulée) détenue par le ménage i : $q_i = \frac{x_i}{n\mu}$, Q_i la proportion de revenu cumulée $Q_i = \sum_{j=1}^{i} q_j$ et P_i la proportion de rang cumulé : $P_i = \frac{\sum_{j=1}^{i} R_{(t_j)}}{n}$. En liant les points de coordonnées (P_i, Q_i) dans le plan (P, Q), on obtient une courbe de concentration. Celle-ci n'est pas forcément monotone croissante. Le coefficient de concentration peut être négatif ou supérieur à 1 si certaines valeurs de x_i sont négatives. En prenant la somme des aires des trapèzes situés en dessous de la courbe de concentration, on obtient une expression équivalente à l'équation (10) :

$$A = \frac{1}{2n}\sum_{i=1}^{n} q_i[1 + 2(n-i)]. \tag{19}$$

Comme l'indice de Gini, l'indice de concentration se mesure par $1 - 2A$:

$$c\left(\frac{x}{t}\right) = 1 - \frac{1}{n}\sum_{i=1}^{n} q_i[1 + 2(n-i)]. \tag{20}$$

Après quelques transformations algébriques, l'indice de concentration peut se réécrire comme suit :

$$c\left(\frac{x}{t}\right) = 2\,\mathrm{Cov}\,(q(x), R_{(t_i)}) = \frac{2}{n\mu}\mathrm{Cov}\,(x, R_{(t_i)}). \tag{21}$$

Cette expression montre que le coefficient de concentration est calculé à partir de la variable x et du critère de classification t. Une équivalence s'établit alors avec la courbe de Lorenz et l'indice de Gini. Il faut pour cela considérer que x_i est le revenu du ménage i et que le rang de chaque ménage correspond au rang après classement croissant des revenus : $R_{(t_i)} = : R_{(x)}$. Étant donné que $t_i = 1$ pour la plus faible valeur de t, alors $R_{(x)} = 1$ pour le plus faible revenu et $R_{(x)} = n$ pour le plus élevé. L'équivalence s'exprime alors par :

$$G = c\left(\frac{x}{x}\right) = 2\,\mathrm{Cov}\left(q(x), R_{(x)}\right) = \frac{2}{n\mu}\mathrm{Cov}\left(x, R_{(x)}\right). \tag{22}$$

L'indice de Gini est un cas particulier du coefficient de concentration et la courbe de Lorenz est un cas particulier de la courbe de concentration. Ces cas particuliers sont construits en considérant que le rang attribué à chaque famille correspond à l'ordre dans lequel elle apparaît après classement croissant des revenus. Le revenu de chaque ménage étant la résultante de q facteurs additifs, le coefficient de Gini s'écrit alors :

$$G = 2\,\mathrm{Cov}\left(x, R_{(x)}\right) = \frac{2}{n\mu}\mathrm{Cov}\left(\sum_{m=1}^{q} x_i^m, R_{(x)}\right) = \sum_{m=1}^{q} \phi_m c\left(\frac{x^m}{x}\right). \tag{23}$$

Il s'agit de l'expression formulée par Rao (1969). Le pseudo-Gini de la source m s'exprime par :

$$\bar{G}_m = c\left(\frac{x^m}{x}\right) = \frac{2}{n\mu^m}\mathrm{Cov}\left(x^m, R_{(x)}\right). \tag{24}$$

Au contraire, l'indice de Gini mesuré sur la source m est :

$$G_m = c\left(\frac{x^m}{x^m}\right) = \frac{2}{n\mu^m}\mathrm{Cov}\left(x^m, R_{(x^m)}\right). \tag{25}$$

À partir des deux équations précédentes, un autre lien permet d'unir l'indice de Gini et son pseudo-Gini. En effet, le i^e point sur la courbe de concentration autorise le calcul de l'indice de Gini associé à la source m qui regroupe tous les revenus inférieurs à i. En prenant l'indice de Gini $c\left(\frac{x^m}{x^m}\right)$, la somme des sources m inférieures à i ne peut pas excéder celle qui est obtenue à partir du pseudo-Gini $c\left(\frac{x^m}{x}\right)$. Par conséquent, la courbe de concentration du pseudo-Gini procure toujours un indice moins important que l'indice de Gini. D'où la relation :

$$c\left(\frac{x^m}{x}\right) = \bar{G}_m \leq G_m \Leftrightarrow \frac{c\left(\frac{x^m}{x}\right)}{G_m} = \Omega_m \leq 1. \tag{26}$$

Autrement dit, les indices de concentration $c\left(\frac{x}{t}\right)$ ne peuvent jamais excéder la valeur de leur indice de Gini.

Pour parvenir à une différentiation des résultats obtenus à partir de données groupées ou de données individuelles, calculons l'indice de Gini à partir de données groupées à la manière de Rao (1969). Les ménages sont regroupés en k classes ($j = 1, ..., k$). On note μ_j la moyenne des revenus de la classe j et $\vec{\mu}$ le vecteur des moyennes de chaque classe. L'indice de Gini mesuré sur les revenus des ménages regroupés en j classes est :

$$G(\vec{\mu}) = c\left(\frac{\vec{\mu}}{\vec{\mu}}\right) = \frac{2}{n\mu}\mathrm{Cov}\left(\vec{\mu}, R_{(\vec{\mu})}\right). \tag{27}$$

D'autre part, la relation suivante est toujours vérifiée :

$$\mathrm{Cov}\ (\vec{\mu}, R_{(\vec{\mu})}) = \mathrm{Cov}\ (x, R_{(\mu_j)}). \tag{28}$$

On sait aussi que : $\mathrm{Cov}\ (x, R_{(\mu_j)}) \leq \mathrm{Cov}\ (x, R_{(x)})$. En conséquence :

$$\mathrm{Cov}\ (x, R_{(\mu_j)}) = \mathrm{Cov}\ (\vec{\mu}, R_{(\vec{\mu})}) \leq \mathrm{Cov}\ (x, R_{(x)}) \Longrightarrow G\ (\vec{\mu}) \leq G\ (\mathrm{x}). \tag{29}$$

En définitive, l'indice de Gini estimé à partir de données en classes sous-estime en permanence l'indice de Gini mesuré à partir de données individuelles. Il est donc important d'utiliser les techniques de décomposition en sources de revenu pour lesquelles les données sont exprimées par individu. Les auteurs étudient néanmoins les conditions dans lesquelles l'erreur peut revêtir une forme particulière selon les rangs considérés. Une difficulté d'interprétation subsiste car la décomposition de l'indicateur de Gini dépend de l'indice de corrélation des rangs Ω_m :

$$G = \sum_{m=1}^{q} \phi_m \Omega_m G_m\ . \tag{30}$$

Selon FRK, Ω_m reste constant dans le temps. En revanche, Pyatt, Chen et Fei estiment que ce coefficient est dépendant d'un aspect dynamique du développement, comme le départ des agriculteurs vers des emplois salariés. Cette idée d'application au domaine agricole sera ensuite reprise par Adams (1994) sur 734 ménages pakistanais.

En définitive, les approches précédentes montrent que les estimations peuvent conduire à trois types d'erreur. La première est l'utilisation de l'équation (7), autrement dit, considérer que G dépend uniquement de G_m ($\forall m$). La deuxième est l'utilisation de données groupées. La troisième est une combinaison des deux précédentes. Par ailleurs, on notera que les techniques de décomposition ont chacune apporté de nouveaux éléments qui ont permis de perfectionner et d'étendre la méthodologie de la décomposition de l'indice de Gini et celle de l'indice de concentration. Une césure va pourtant s'effectuer au sein de la littérature puisqu'en 1982, Shorrocks va exprimer des doutes concernant la décomposition de la mesure de Gini.

3. L'apport axiomatique de Shorrocks (1982)

En 1982, Shorrocks propose de faire une synthèse des travaux effectués dans le domaine de la décomposition des mesures d'inégalité en sources de revenu. Son approche est qualifiée d'analyse axiomatique. Il refuse cependant cette appellation. Il introduit différentes règles de décomposition sous forme d'hypothèses, puis propose en 1983 une application aux revenus américains.

Hypothèse 1 : Symétrie, continuité et normalisation.

La mesure d'inégalité est une fonction continue et symétrique de ses arguments[5]. La normalisation est considérée comme une borne inférieure. Si tous les revenus sont égaux à $\varepsilon \in \mathbb{R}_+$, alors : $I(\varepsilon, ..., \varepsilon) = 0$.

Hypothèse 2 : (a) Continuité et (b) traitement symétrique des facteurs.

(a) La continuité est définie par : $S^m(x; q)$ est continue en x^m où S^m est la contribution de la source m à l'inégalité totale, x le vecteur des revenus et q le nombre de facteurs.

(b) Le traitement symétrique des facteurs est :

$$S^m(x^1, x^2, ..., x^m; q) = S^{\Pi m}(x^{\Pi 1}, x^{\Pi 2}, ..., x^{\Pi m}; q)$$

où Πm est une permutation des éléments du vecteur des sources m. Cela signifie, que l'ordre des éléments à l'intérieur de la distribution x^m ne modifie pas la valeur de la contribution $S^m(x; q)$.

Hypothèse 3 : Indépendance par rapport au niveau de désagrégation.

La contribution $S^m(x; q)$ ne doit pas dépendre de la manière dont les revenus sont distingués ou décomposés. Par exemple, si le revenu est divisé en salaires, primes, rentes et transferts ou seulement en salaires et autres revenus, alors la contribution du facteur salaire à l'inégalité totale ne doit pas être modifiée[6]. Autrement dit, si la

5. En effet, lorsqu'une perturbation est introduite dans la distribution, la mesure d'inégalité reste inchangée. La symétrie équivaut aussi au principe qui permet de conserver l'anonymat des individus.

6. Lorsqu'une décomposition en facteurs est réalisée par la valeur de Shapley (1953), on ne respecte pas cette propriété (*Cf.* section 6 *infra* sur les méthodes permettant de généraliser les techniques de décomposition). Il est nécessaire de recourir à d'autres valeurs comme celle d'Owen (1977).

contribution de la source 1 est évaluée conjointement avec quatre autres sources ou par rapport à un ensemble qui regroupe ces quatre sources, alors la contribution reste inchangée :

$$S^1(x^1, x^2, ..., x^4; q) = S^1\,(x^1, x - x^1; 2) = S^1\,(x^1, x).$$

Hypothèse 4 : Décomposition cohérente.

Elle implique trivialement que la somme des contributions $S^m(x;\, q)$ procure la valeur globale de l'inégalité $I\,(x)$. Elle est définie par :

$$\sum_{m=1}^{q} S^m(x^1, x^2, ..., x^m, ..., x^q; q) = \sum_{m=1}^{q} S^m(x^m; q) = I(x)\,. \tag{31}$$

Ces quatre hypothèses de base permettent de spécifier un théorème dont le résultat montre qu'une mesure d'inégalité se réécrit comme une moyenne pondérée des sources de revenu. En effet, pour une population de n individus :

$$I(x) = \sum_{m=1}^{q} S^m(x^m; q) = \sum_{i=1}^{n}\sum_{m=1}^{q} a_i x^m = a \cdot \sum_{m=1}^{q} x^m = ax\,. \tag{32}$$

En corollaire, la contribution de la source m (S^m) à l'inégalité totale se formule avec les mêmes poids. Par exemple, pour l'indicateur de Gini, on a :

$$S^m(x^m, x) = \frac{2}{n^2\mu}\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)x_i^m = \frac{\mu^m}{\mu}\,\bar{G}(x^m)\;^{7}. \tag{33}$$

Hypothèse 5 : (a) Symétrie de la population et (b) normalisation des distributions de facteurs équivalents.

(a) Les contributions de chaque facteur à l'inégalité totale ne doivent pas dépendre de la manière avec laquelle les individus sont classés : $S^m\,(x^m\Pi, x\Pi) = S^m\,(x^m, x)$. En introduisant une perturbation dans la distribution globale (par une matrice de permutation qui change l'ordre des individus) et une perturbation dans la distribution de la source m, le rang des individus est modifié. Cette perturbation ne doit pas affecter la valeur de la contribution. On constate immédiatement que cette restriction invalide la décomposition de l'indice de Gini, car le pseudo-Gini dépend du rang de la distribution globale.

(b) La normalisation pour des distributions de sources équivalentes permet d'attribuer la valeur nulle à la contribution de la source m lorsque le vecteur de la source m est constitué d'un même réel. Par exemple, si le vecteur est unique-

7. Rappelons que cette décomposition est valable lorsque les revenus sont ordonnés de manière croissante et où i représente le rang de l'individu i.

ment constitué de la valeur moyenne μ^m, alors la contribution de la source m à l'inégalité totale est nulle : $S^m(\mu^m \mathbb{1}, x) = 0, \forall \mu^m$, où $\mathbb{1}$ est le vecteur « unité » (1, ..., 1) de taille n. Cette hypothèse est nécessaire mais non suffisante. Elle permet de définir une décomposition unique pour chaque indicateur. Le problème est que l'unicité est seulement valable pour une population de deux personnes. Les décompositions ne sont donc pas naturelles[8]. Il faut alors introduire une hypothèse plus restrictive.

Hypothèse 6 : Symétrie de deux facteurs.

Il s'agit d'une restriction de l'hypothèse 2 (b) pour deux facteurs : x^1 et x^2. Supposons que x^2 soit une permutation de x^1 : $x^2 = x^1\Pi$, où Π est une matrice de permutation de taille $n \times n$. Cette hypothèse permet de définir l'égalité entre la contribution issue de x^1 et son permuté x^2. Lorsqu'il existe plusieurs variables, la corrélation entre x^1 et $x - x^1$ n'est pas la même que celle qui prévaut entre x^2 et $x - x^2$. Ces différences de corrélation donnent des contributions à l'inégalité totale différentes. En revanche, lorsqu'il existe seulement deux facteurs, ces contributions doivent être égales : $S^1(x^1, x^1 + x^1\Pi) = S^2(x^1\Pi, x^1 + x^1\Pi)$.

Théorème 3.1 *Les hypothèses 1 – 6 impliquent :*

$$s^m(I) = \frac{S^m(x^m, x)}{I(x)} = \frac{\mathrm{Cov}(x^m, x)}{\sigma^2(x)}. \tag{34}$$

Ce théorème fournit trois résultats primordiaux. Premièrement, la règle de décomposition est unique. Deuxièmement, la contribution relative (s^m) de chaque source à l'inégalité totale est indépendante du choix de la mesure. Cela signifie que pour l'indice de Gini, s'il existe une équation qui satisfait les hypothèses 1-6, la contribution relative pourrait alors se mesurer avec la covariance et la variance ($\sigma^2(x)$). Troisièmement, la variance et le coefficient de variation vérifient les hypothèses 1 à 6. Ceci permet d'invalider, en partie, toutes les recherches qui ont été menées pour démontrer la pertinence de l'utilisation de la décomposition de l'indicateur de Gini. En effet, en reprenant les axiomes introduits par Shorrocks (1982), il est possible de vérifier que la décomposition de la mesure de Gini proposée par Rao (1969 : éq. (4)), celle de FRK (1978 : éq. (12)), celle de Fields (1979 : éq. (16)), celle de Pyatt, Chen Fei (1980 : éq. (24)) ou encore celle proposée par Shorrocks (1982 : éq. (34)) ne satisfont pas l'hypothèse de traitement symétrique des facteurs (H2 (b)), l'hypothèse de symétrie de la population (H5 (a)) et l'hypothèse de symétrie de deux facteurs (H6). En effet, étant donné que la mesure de Gini décomposée dépend du classement de chaque ménage selon leur niveau de vie, les axiomes d'indépendance par rapport aux rangs ne peuvent être respectés.

8. Le terme naturel renvoie à l'unicité de la décomposition.

Notons néanmoins, que les hypothèses introduites pas Shorrocks ne font pas office de référence puisque certains auteurs vont proposer des principes contradictoires (*Cf.* par exemple Morduch et Sicular, 2002 : section 5.1 *infra*). D'autres chercheurs, vont par ailleurs continuer à démontrer la pertinence des mesures qui dépendent du rang des individus afin de nuancer les propos de Shorrocks (1982). Lerman et Yitzhaki (1985) vont par exemple soutenir que la décomposition de la mesure de Gini reste « désirable ».

4. La décomposition de l'indice de Gini étendu : Lerman et Yitzhaki (1985)

En partant de la formule de l'indice de Gini basée sur la covariance, notamment développée et généralisée par Pyatt, Chen et Fei (1980), Lerman et Yitzhaki (1985) présentent la décomposition de l'indice de Gini étendu. Il s'agit de l'indicateur de Gini qui tient compte de l'aversion de la collectivité à l'égard des inégalités. En d'autres termes, il intègre la préférence relative pour l'égalité des revenus.

On appelle F la fonction de réparation des revenus et F^m la fonction de répartition du facteur m. Les distributions de revenus sont uniformément réparties entre [0,1] afin que la moyenne soit ½. L'indicateur de Gini se formule par $G = \frac{2\,\mathrm{Cov}(x,F)}{\mu}$. Il peut se réécrire en tenant compte de l'ensemble des sources de revenu :

$$G = \frac{2\sum_{m=1}^{q}\mathrm{Cov}(x^m,F)}{\mu}. \tag{35}$$

Afin de faire apparaître l'indice de Gini de la source m, il est possible de multiplier et de diviser l'expression précédente par des termes qui se neutralisent :

$$G = \frac{2\sum_{m=1}^{q}\mathrm{Cov}(x^m,F)}{\mu} = \sum_{m=1}^{q}\frac{\mathrm{Cov}(x^m,F)}{\mathrm{Cov}(x^m,F^m)}\cdot\frac{2\,\mathrm{Cov}(x^m,F^m)}{\mu^m}\cdot\frac{\mu^m}{\mu}. \tag{36}$$

On obtient alors la relation suivante :

$$G = \sum_{m=1}^{q} R_m G_m \phi_m \tag{37}$$

où

$$R_m = \frac{\mathrm{Cov}(x^m,F)}{\mathrm{Cov}(x^m,F^m)},\ G_m = \frac{2\,\mathrm{Cov}(x^m,F^m)}{\mu^m},\ \phi_m = \frac{\mu^m}{\mu}. \tag{38}$$

R_m est définie comme la corrélation de Gini entre la source m et le revenu global (équivalent à Ω_m dans l'équation (15) introduite par FRK, 1978 et reprise par Fields, 1979). Rappelons que : G_m est l'indice de Gini associé à la source m, ϕ_m est la proportion du facteur m dans le revenu moyen et que R_m est similaire au coefficient de corrélation de Pearson. En effet, $R_m \in [-1, 1]$.

- Si $R_m = 1$, la source m est une fonction croissante du revenu total.
- Si $R^m = -1$, le facteur m est une fonction décroissante du revenu total.
- Si $R_m = 0$, la source de revenu m est également distribuée et sa contribution à l'inégalité totale est nulle[9].

Ceci nous conduit à l'évaluation de l'impact marginal de la source m sur l'indice global. Cette mesure s'effectue par une simple dérivée partielle. Supposons que chaque personne voit le montant de sa source de revenu m varier de εx^m, où $\varepsilon x^m \to 1$. Il est donc possible de mesurer la dérivée partielle du Gini global compte tenu d'une petite variation en pourcentage de la source m (représentée par ε^m) :

$$\frac{\partial G}{\partial \varepsilon^m} G = \phi^m (R_m G_m - G)\,. \tag{39}$$

D'après Yitzhaki (1983), il est possible de construire un indice de Gini qui dépend de l'intensité de la préférence sociale pour l'égalité. Si cette intensité est captée par le paramètre ν, alors[10] :

$$G(\nu) = 1 - \nu\,(\nu - 1)\int_0^1 (1-F)^{\nu-2} L(F)\, dF \tag{40}$$

où $G\,(\nu)$ est l'indice de Gini étendu (ou généralisé) et où $L\,(F)$ est la courbe de Lorenz. Le paramètre de préférence pour l'égalité est compris entre 0 et l'infini. Lorsque :

- $\nu \in [0, 1[$, l'indice reflète la préférence pour l'inégalité;
- $\nu = 1$, l'indice reflète l'indifférence à l'égalité;
- $\nu = 2$, le coefficient de Gini généralisé est égal à l'indice de Gini standard;
- $\nu \to \infty$, le critère de Rawls est atteint, autrement dit, la parfaite égalité est préférée.

La décomposition en sources de revenu est analogue à la précédente :

---

9. Ce dernier point est discutable puisque des sources également distribuées peuvent réduire l'inégalité relative (*Cf.* Morduch et Sicular, 2002 : section 6.1, et Araar, 2006).

10. En théorie de la décision, l'indice de Gini généralisé procure par dualité une fonction de bien-être social dans laquelle une fonction associée à ν représente le comportement du preneur de décision qui évolue dans un univers incertain ou risqué (*Cf.* Chateauneuf, Gajdos, et Wilthien, 2002).

$$G = \sum_{m=1}^{q} R_m(\nu)\, G_m(\nu)\, \phi_m \tag{41}$$

où

$$R_m(\nu) = \frac{\mathrm{Cov}\left(x^m, (1-F^0)^{\nu-1}\right)}{\mathrm{Cov}\left(x^m, (1-F^m)^{\nu-1}\right)},\ G_m(\nu) = \frac{-\nu\,\mathrm{Cov}\left(x^m, (1-F^m)^{\nu-1}\right)}{\mu^m}. \tag{42}$$

Lerman et Yitzhaki (1985) s'interrogent sur la cohérence de leur décomposition. Pour cela, ils s'appuient sur les hypothèses introduites par Shorrocks (1982). La décomposition de Lerman et Yitzhaki tombe dans la catégorie des multiples décompositions de l'indicateur de Gini, dont la séparation n'est pas unique. Selon Shorrocks, l'idée de choisir l'indice de Gini est « non acceptable ». Selon Lerman et Yitzhaki, leur décomposition est « désirable ». Ils font ici allusion au fait que l'indice de Gini est compatible avec le concept de dominance stochastique et les principes de transferts qui y sont sous-jacents[11]. Cette décomposition est ensuite beaucoup plus attrayante du fait qu'en plus de réunir les trois indices (l'indice de Gini de chaque facteur, la part de chaque source dans le revenu moyen et la corrélation de la source avec le rang des individus), elle prend en compte la préférence pour l'égalité des agents. L'approche de Shorrocks est une analyse mathématique très intéressante, mais le fait que les poids (a_i) (*Cf.* éq. (32)) procurent de multiples décompositions n'est pas, selon les auteurs, explicité de manière économique. Enfin, la dernière idée qui incite à privilégier cette décomposition est le calcul des contributions marginales. Il s'agit de mesurer la variation d'une source précise m et d'évaluer son impact sur l'indice global. L'application de cette méthode révèle par exemple que la contribution marginale des revenus féminins (dans un ménage) excède celle issue du revenu du capital[12].

En somme, on remarquera que cette nouvelle décomposition en sources de revenu minimise l'influence exercée par l'indice du pseudo-Gini[13].

4.1 *Une application à la consommation : Garner (1993)*

L'approche de Lerman et Yitzhaki (1985) est intéressante, car l'analyse des contributions marginales constitue un apport fondamental. Les contributions marginales peuvent être évaluées par de petites variations inhérentes aux valeurs d'une source de revenu. Supposons qu'il existe un changement dans les dépenses

11. *Cf.* Makdissi et Mussard (2006).

12. Notons que Shorrocks (1983) atteint le même résultat sans utiliser le concept de contribution marginale. En 1989, Lerman et Yitzhaki utilisent cette méthode pour étudier les inégalités de sources de revenu des ménages américains de 1983.

13. Lerman et Yitzhaki (1985) dérivent leur mesure sans mentionner la mesure du pseudo-Gini. Néanmoins, il est possible de remarquer l'étroite relation avec l'approche de Rao (1969), FRK (1978), Fields (1979) et celle de Pyatt, Chen et Fei (1980).

de consommation. Ce changement concerne une source m. L'intensité du changement est mesurée par le pourcentage ε^m. On considère ensuite que cette variation concerne l'ensemble des individus d'une population sur laquelle sont mesurées les inégalités. Ce changement marginal peut être évalué en proportion de la valeur de l'indice de Gini global. Il est appelé contribution marginale relative :

$$\frac{\frac{\partial G}{\partial \varepsilon_m}}{G} = \underbrace{\frac{R_m G_m \phi_m}{G}}_{s^m} - \phi_m \,. \tag{43}$$

Cette formule est très attrayante. Contrairement à Lerman et Yitzhaki (1985), Garner mesure le poids relatif de l'impact marginal. Ceci aboutit à une définition claire. Il s'agit de la contribution relative des inégalités de la source m (s^m) à laquelle on retranche la contribution moyenne de la dépense de consommation du facteur m. Cela permet d'envisager et de mesurer l'effet d'une taxe imposée sur une source particulière. Lorsque la contribution marginale s^m est positive, une taxe sur ce facteur permet de réduire l'inégalité totale. Cette taxe doit être progressive afin d'affecter marginalement les personnes riches en majorité. En revanche, lorsque la contribution relative s^m est négative, une taxe sur le bien de consommation a pour conséquence d'augmenter l'inégalité totale. En outre, ces résultats doivent tenir compte de la nature des biens consommés.

D'après Yitzhaki (1990), il est possible de calculer l'élasticité de la source m par rapport à la consommation globale :

$$\eta^m = \frac{R_m G_m}{G} = b^m \frac{\mu}{\mu^m} \tag{44}$$

où b^m peut être considéré comme un paramètre de régression ou comme la propension moyenne à consommer le bien m. Les valeurs des élasticités renvoient à des biens de consommation particuliers :

- $\eta^m > 1$ caractérise les biens de luxe;
- $\eta^m \in [0, 1]$ caractérise un bien de première nécessité;
- $\eta^m < 0$ caractérise un bien inférieur.

Une application est menée sur les dépenses de consommation américaines de 1987. L'indice de Gini global de 0,33 est décomposé[14]. Les élasticités sont notamment très élevées pour les dépenses vestimentaires et les loisirs. Une conséquence de ces résultats est que la taxation de ces biens augmenterait la progressivité du système fiscal[15].

14. La valeur particulière $G = 0{,}33$ est analysée par Milanovic (1997).

15. *Cf.* Araar (2002) pour l'impact de la variation des prix sur l'inégalité et le bien-être.

4.2 *Une application à la redistribution : Podder (1993)*

Une autre élasticité est ici considérée. Elle mesure l'effet de l'augmentation de 1 % de la moyenne de la source m sur l'inégalité totale.

Théorème 4.1 *Soit une variation de la moyenne de la source m (μ^m) sans changement de sa courbe de concentration. L'élasticité de l'indice de Gini par rapport à sa source m est donnée par :*

$$\eta^m = \frac{1}{G}\left[\frac{\mu^m}{\mu}(c_m - G)\right] \tag{45}$$

où c_m est l'indice de concentration associé à la source de revenu m.

Il est possible de démontrer que cette élasticité est égale à :

$$\eta^m = \frac{\frac{dG}{G}}{\frac{d\mu^m}{\mu^m}}, \quad \sum_{m=1}^{q} \eta^m = 0. \tag{46}$$

La nullité de la somme des élasticités implique qu'un changement proportionnel concernant toutes les sources m ne modifie pas l'indice de Gini. Les implications de la méthode ne concernent pas la fiscalité. Elles concernent les effets de la redistribution. En effet, la redistribution d'une source particulière (notamment une augmentation de x % de la source m) peut affecter la répartition globale des revenus et par conséquent modifier les inégalités globales. Cet outil permet donc de mesurer l'impact d'une mesure de redistribution (pour une source m) concernant l'ensemble de la population.

L'approche de Lerman et Yitzhaki peut être considérée comme fondamentale. Elle a permis de créer une approche en terme d'élasticité de Gini, utilisée à la fois en fiscalité et en redistribution tout en permettant d'apprécier l'impact du travail des femmes sur l'évolution du marché du travail et sur celui des distributions de revenu (*Cf.* Cancian et Reed, 1998)[16].

5. Les tentatives de généralisation

Les implications de l'approche de Lerman et Yitzhaki (1985) incitent à penser qu'il s'agit de la méthode de décomposition la plus élaborée. Pourtant, les recherches en ce domaine ne restent pas vaines. En effet, les développements récents s'attachent à généraliser la technique de décomposition en sources de revenu en privilégiant deux orientations : l'analyse économétrique et la valeur de Shapley (1953).

16. *Cf.* Flückiger et Silber (1995) dont la méthode permet d'expliquer la différence entre l'indicateur de Gini par facteur des hommes et celui des femmes.

5.1 *L'approche économétrique : Morduch et Sicular (2002)*

Les sections précédentes montrent que le chercheur peut être contraint par la structure additive qu'implique le recours aux sources de revenu. Une solution peut consister à privilégier l'approche économétrique. Elle paraît à première vue plus générale, car les facteurs peuvent être étudiés sans qu'il n'existe *a priori* de structure additive permettant de les lier. Il est par conséquent concevable d'utiliser des facteurs qui ne sont pas des sources de revenu, autrement dit, des attributs (dimensions) tels que l'éducation, la santé, *etc*., qui peuvent expliquer le niveau de revenu.

Les auteurs critiquent les approches traditionnelles de décomposition en remarquant l'absence de contrôle vis-à-vis de l'endogénéité des variables utilisées. En effet, les sources sont issues des revenus, eux-mêmes déterminés par d'autres variables. Ces problèmes ont incité les chercheurs à examiner la piste économétrique. Plus précisément, Morduch et Sicular introduisent la régression dans la décomposition. Ce recours permet notamment de contrôler l'endogénéité des variables et d'intégrer leur continuité. La méthode permet de calculer : *(i)* l'exacte contribution d'une source particulière à l'inégalité totale, *(ii)* son écart-type, *(iii)* son intervalle de confiance. Pour cela, deux définitions de base sont proposées.

Définition 5.1 ***Additions uniformes*** *: Un indice d'inégalité I(x) décroît lorsque les individus reçoivent un même transfert positif t :*

$$I\,(x + t\mathbb{1}) \leq I\,(x),\ t \in \mathbb{R}_+, \tag{47}$$

où $\mathbb{1}$ est le vecteur unité.

Il s'agit d'une propriété issue de l'axiome de transfert de Pigou-Dalton. La mesure d'inégalité décroît lorsqu'une personne riche transfère une partie de son revenu à une personne pauvre. Rappelons que x^m est le vecteur représentant la m^e source de revenu des n individus. La définition 1 s'exprime alors par :

Définition 5.2 *Une méthode de décomposition qui attribue à chaque facteur une proportion d'inégalité s^m, satisfait la propriété d'additions uniformes si pour tout vecteur $x^m = t\mathbb{1}$, alors $s^m < 0$.*

Autrement dit, toute méthode de décomposition qui accorde une contribution relative négative à une source de revenu (pour des distributions de facteur égalitaires) satisfait le principe d'additions uniformes. Néanmoins la définition 5.1 n'implique pas la définition 5.2. Par exemple, l'indice de Gini satisfait la définition 5.1, mais sa décomposition en pseudo-Gini ne permet pas de vérifier la définition 5.2 [17].

17. De même, la décomposition du coefficient de variation (σ / μ) en sources de revenu aboutit au rejet de la définition 5.2. À ce stade, on constatera aussi que la définition 5.2 remet en cause les anciennes méthodes de décomposition de la mesure de Gini. *Cf.* Araar (2006) pour ce type de problème.

Soit un modèle de régression linéaire multiple de type :

$$x = Y\beta + \varepsilon. \tag{48}$$

La matrice Y comprend en première colonne le vecteur $\mathbb{1}$, puis les autres colonnes sont constituées par la matrice X. La matrice Y est donc de taille $n \times (q + 1)$. L'estimation est faite par moindres carrés (régression quantile) en tenant compte de différentes corrections (endogénéité, *etc*.). Cette méthode est donc plus générale. Elle peut en effet être utilisée pour des facteurs de revenu ou bien attributs (éducation, âge, santé, *etc*.). Par conséquent, le revenu estimé en fonction des attributs est :

$$\hat{x} = Y\hat{\beta}. \tag{49}$$

De même, les revenus issus de l'attribut m sont :

$$\hat{x}^m = Y\hat{\beta}^m. \tag{50}$$

Le revenu global de l'individu i est : $x_i = \sum_{m=1}^{q+1} \hat{x}_i^m$ où $\hat{x}_i^m = \hat{\varepsilon}_i$ si $m = q + 1$. Rappelons que d'après Shorrocks (1982), les contributions relatives de chaque source se formulent par :

$$s^m = \frac{\sum_{i=1}^{n} a_i(x_i^m)}{I(x)}. \tag{51}$$

Ainsi, la contribution relative d'un facteur à l'inégalité totale devient :

$$s^m = \hat{\beta}^m \left(\frac{\sum_{i=1}^{n} a_i(x_i^m)}{I(x)} \right). \tag{52}$$

Il devient alors possible de décomposer tous les indices d'inégalité qui peuvent s'écrire comme une moyenne pondérée des revenus. D'autre part, peu d'études mettent l'accent sur les erreurs liées à l'estimation des inégalités. Pourtant, étant donné la linéarité dans l'estimation des paramètres, on a :

$$\sigma(s^m) = \sigma(\hat{\beta}^m) \left(\frac{\sum_{i=1}^{n} a_i(x_i^m)}{I(x)} \right). \tag{53}$$

L'écart-type des erreurs s'exprime par :

$$\sigma(s^\varepsilon) = \left\{ \sigma_\varepsilon^2 \sum_{i=1}^{n} \left(\frac{a_i(x)}{I(x)} \right)^2 \right\}^{\frac{1}{2}}. \tag{54}$$

En conséquence, l'expression (54) permet le calcul des intervalles de confiance des contributions relatives sous l'hypothèse d'homoscédasticité.

En définitive, on remarque que cette approche remet en cause les décompositions analytiques (non paramétriques) des inégalités. La première remise en cause s'effectue par les définitions 5.1 et 5.2. La seconde attire l'attention du chercheur sur la contrainte qui stipule que la somme des facteurs doit être égale au revenu. Cependant, l'approche de Morduch et Sicular (2002) permet une certaine généralisation puisqu'il est permis d'utiliser à la fois la méthode pour des sources de revenu ou pour des attributs. Il faut tout de même noter que cette généralisation implique certaines restrictions. Les variables doivent être indépendantes, le résidu doit suivre une loi normale et l'ajustement du modèle doit être de qualité. D'autre part, les auteurs s'appuient sur un socle axiomatique assez restreint en mettant en exergue la propriété d'additions uniformes qui permet de privilégier la mesure de Theil et de rejeter celle de Gini. En conséquence, la méthodologie souffre de ne pas offrir une règle unanime.

5.2 *La valeur de Shapley (1953) : un outil de décomposition et de généralisation*

En 1992, une nouvelle approche originale est apparue dans la littérature. Auvray et Trannoy (1992) proposent de décomposer les mesures d'inégalité par la valeur de Shapley (1953). La généralisation provient du fait que cette technique peut être intégrée à une grande majorité d'indicateurs (les inégalités, la pauvreté, le R^2 d'un modèle de régression, *etc*.). Il s'agit donc d'une avancée radicale au sein de la littérature et certains auteurs en sont convaincus (*Cf.* Silber, 2004). Il s'agit d'une synthèse entre la théorie des jeux coopératifs et l'économie des inégalités. L'utilisation de la valeur de Shapley suggère que l'inégalité globale peut être partagée au même titre que le problème de partage d'une somme d'argent que l'on rencontre en théorie des jeux coopératifs[18].

Au regard des principes de décomposition en sources de revenu introduits par Shorrocks (1982), Chantreuil et Trannoy (1999) indiquent que la décomposition par la valeur de Shapley ignore deux règles cruciales : l'indépendance par rapport au niveau de désagrégation et la normalisation des distributions de facteurs équivalents. Afin de surmonter la difficulté liée à l'indépendance par rapport au niveau de désagrégation, Chantreuil et Trannoy (1999) recommandent l'utilisation de la méthode « *Nested Shapley* ». Shorrocks (1999) va lui-même s'intéresser à cet outil de décomposition en proposant une décomposition par la valeur d'Owen (1977).

Shorrocks identifie tout d'abord un problème de symétrie. Selon lui, la procédure de Shapley envisage l'ensemble de toutes les éliminations possibles des variables considérées. En éliminant les facteurs selon une règle précise, on met

18. *Cf.* Araar (2006) pour une critique complète de la méthode.

en évidence l'impact marginal de chaque facteur. En prenant la moyenne des ces impacts marginaux, on mesure la contribution du facteur à l'indicateur global. Mais dans ce processus, il existe un problème de dépendance, car les facteurs ne sont pas traités de manière symétrique. Par conséquent, un algorithme est proposé afin d'équilibrer le nombre d'impacts marginaux.

Supposons qu'un indice $I(x)$ (Gini, pauvreté, R^2, *etc.*) soit déterminé par q facteurs x^m, $m \in K = \{1, 2, ..., q\}$. La fonction $F(S)$ procure la valeur de l'indice I lorsque le facteur x^m $(m \notin S)$ a été éliminé de l'ensemble des facteurs. On appelle s le nombre de facteurs restant après la séquence d'élimination des x^m : $s := |S|$. La fonction F, est définie par :

$$F : \{S \mid S \subseteq K\} \longrightarrow \mathbb{R}. \tag{55}$$

Les q! éliminations possibles des q facteurs permettent de mesurer la contribution d'un facteur m à l'inégalité totale. On mesure ainsi la contribution du facteur considéré quand la séquence d'élimination est faite au hasard. Les facteurs sont donc traités de manière symétrique. En conséquence, la contribution du facteur m à la valeur totale de l'indice $I(x)$ est donnée par :

$$C_m^s(K,F) = \sum_{s=0}^{q-1} \sum_{S \subseteq K \setminus \{m\}} \frac{(q-1-s)!s!}{q!} \Delta_m F(S) \tag{56}$$

où $\Delta_m F(S)$ est l'impact marginal du facteur m lorsque ce dernier a été retiré de l'ensemble des facteurs S :

$$\Delta_m F(S) := F(S \cup \{m\}) - F(S),\ F(\emptyset) = 0. \tag{57}$$

La contribution (C_m^s) du facteur m à l'indice $I(x)$ se mesure donc en prenant la moyenne des impacts marginaux du facteur m lorsque le processus d'élimination des variables intègre toutes les combinaisons possibles.

Pour démontrer la portée de la méthode de Shapley, Shorrocks (1999) propose une application au modèle de Datt et Ravallion (1992) selon lequel les effets de la croissance économique et de la redistribution expliquent les changements de la pauvreté au cours du temps. Le modèle s'exprime par :

$$\Delta P := F(G, R) \tag{58}$$

où ΔP représente la variation de la pauvreté, G la croissance économique et R la redistribution. Pour mesurer les effets respectifs de G et R sur la variation du niveau de pauvreté, il est possible de mesurer les contributions de chaque facteur en utilisant la valeur de Shapley. Les contributions de la croissance et de la redistribution sont données par :

$$C_G^s = \frac{1}{2} \left(F(G, R) - F(R) + F(G)\right) \tag{59}$$

$$C_R^s = \frac{1}{2}\left(F(G,R) - F(G) + F(R)\right). \tag{60}$$

Étant donné que la somme de ces deux composantes est égale à la valeur de la variation de la pauvreté ΔP, on peut exactement déterminer quelle est la contribution (s^m) de chaque facteur. Shorrocks généralise ensuite ce résultat aux mesures de pauvreté, notamment à la classe des mesures de Foster-Greer-Thorbecke (1984).

Contrairement à ce que préconisent Chantreuil et Trannoy (1999), Shorrocks (1999) s'interroge sur la possibilité de recourir à la valeur d'Owen (1977). Elle permettrait de traiter les structures hiérarchiques au sein des sources de revenu. La première partition des facteurs est l'ensemble $\mathcal{A}$. Cet ensemble de q facteurs primaires varie avec le processus d'élimination des variables. Lorsqu'une ou plusieurs variables sont éliminées, l'ensemble $\mathcal{A}$ devient l'ensemble $\mathcal{T}$ (où $t := |\mathcal{T}|$ est le nombre de facteurs primaires restants après les différentes éliminations). La première phase consiste à définir, par la valeur de Shapley, la contribution du facteur L ($L \in \mathcal{A} = \{1, 2, ..., q\}$) :

$$C_L^O(K,\mathcal{A},F) = \sum_{t=0}^{q-1} \sum_{T \subseteq \mathcal{A}\setminus\{L\}} \frac{(q-1-t)!\,t!}{q!} \Delta_L F(T) \tag{61}$$

où

$$\Delta_L F(T) := F(T \cup \{L\}) - F(T),\ F(\emptyset) = 0. \tag{62}$$

La deuxième phase consiste à appliquer l'algorithme de Shapley à la contribution de la variable L à I (pour tout L), L formant un sous-ensemble de la première partition $\mathcal{A}$. Cette seconde partition L comprends z facteurs. Lorsqu'un ou plusieurs facteurs sont éliminés, l'ensemble L est défini par l'ensemble S ($s := |S|$ étant son cardinal, autrement dit, le nombre de facteurs restant après les éliminations possibles). Par conséquent, la contribution du facteur m à la contribution C_L^O préalablement définie, est :

$$C_{m,L}^O(K,\mathcal{A},F) = \sum_{s=0}^{z-1} \sum_{S \subseteq L\setminus\{m\}} \frac{(z-1-s)!\,z!}{z!} \Delta_m F(S). \tag{63}$$

Si les facteurs m de l'ensemble L sont eux-mêmes fonctions d'une troisième partition de variables, la procédure se poursuit jusqu'à ce qu'il n'y ait plus de partitions. Cette technique en multiniveaux respecte toujours la cohérence agrégative (quel que soit le nombre de partitions). En effet, la somme des contributions donne bien l'indice I :

$$\sum_{L=1}^{q} \sum_{m=1}^{L} C_{m,L}^O = I(x). \tag{64}$$

Le problème que l'on peut rencontrer avec la valeur de Shapley est le suivant. Supposons que l'on applique la décomposition par Shapley au dernier niveau de partitions (autrement dit à tous les facteurs) afin de mesurer la contribution d'un facteur de la première partition. Le résultat sera différent de celui de la décomposition appliquée au premier niveau de la partition. C'est pourquoi, la procédure d'Owen en plusieurs niveaux est recommandée. La méthode *Nested Shapley* est aussi appropriée (*Cf.* Chantreuil et Trannoy, 1999)[19].

Conclusion

La décomposition de la mesure de Gini en facteurs permet de comprendre qu'une source particulière puisse être à l'origine des inégalités de revenu. Cette méthode autorise aussi une analyse complexe des sources de revenu en étroite relation avec différentes thèses de la théorie économique. Par exemple, Fei, Ranis et Kuo (1978) établissent des liens avec la croissance économique et l'économie du développement. La détermination du rôle des sources de revenu dans l'inégalité globale permet enfin d'analyser la distance qui sépare les familles riches des familles pauvres. L'approche de FRK (1978), basée sur les travaux précurseurs de Rao (1969), notamment fondée sur le concept de pseudo-Gini, est le point de départ d'une multitude d'études empiriques. On recense par exemple, des applications sur les sources de revenu colombiennes (Fields, 1979), américaines (Shorrocks, 1983) ou encore pakistanaises (*Cf.* Adams, 1994). Shorrocks (1982) va ensuite généraliser cette décomposition en facteurs en essayant de comprendre pourquoi certains indicateurs, au même titre que la variance et le coefficient de variation, ne sont pas parfaitement décomposables. En somme, une mesure est décomposable en facteurs si elle respecte six propriétés fondamentales. Dans le cas contraire, la décomposition est jugée « non acceptable ». Lerman et Yitzhaki (1985) refusent de respecter ces propriétés et montrent que l'indice de Gini satisfait un procédé de décomposition « désirable » puisqu'il fait apparaître la préférence sociale à l'égard de l'égalité ainsi que les trois indices fondamentaux que sont : la mesure de Gini pour chaque facteur, la part du facteur dans le revenu global et un coefficient de corrélation entre le facteur et le rang des individus selon le revenu global. Une partie de la littérature abandonne donc les six postulats introduits par Shorrocks (1982), puisque les décompositions en facteurs tendent à être unifiées et généralisées, notamment par les techniques économétriques (*Cf.* Morduch et Sicular, 2002) ou les approches axiomatiques telles que la valeur de Shapley (*Cf.* Chantreuil et Trannoy, 1999; Shorrocks, 1999).

En définitive, la lecture de la littérature montre que de nombreux chercheurs, outre Paréto et Keynes, plaident en faveur de la réduction des inégalités dans la

19. Néanmoins, les critiques formulées par Araar (2006) concernant la méthode analytique de décomposition de la mesure de Gini et la décomposition par la valeur de Shapley montrent que les deux méthodes convergent si l'ordre des ménages par rapport à la source m est le même que celui par rapport au revenu global ($\bar{G}^m = G^m$). Lorsque l'interaction entre les sources est importante, la méthode de décomposition par la valeur de Shapley n'est pas appropriée.

répartition des revenus et celle des fortunes, et que le procédé de décomposition de la mesure de Gini en sources de revenu continue à intéresser de nombreux chercheurs, qui essaient notamment d'unifier cette méthode et la décomposition en sous-groupes. Rao (1969) proposa, de manière exclusive, les méthodes de décomposition en sous-groupes et en sources de revenu. Pourtant, les deux techniques sont passibles d'une synthèse, la multidécomposition, aboutissant à la construction de nouveaux indices décomposés selon deux dimensions : les groupes et les facteurs (*Cf*. Mussard, 2004a, 2004b, 2006). Il est ainsi permis de penser que les techniques de décomposition qui, à l'heure actuelle sont en perpétuelle évolution, permettront d'aboutir à des applications complexes, mais novatrices, permettant par exemple de démontrer que la baisse des inégalités autorise celle des inéquités. Témoignent de l'existence de cette perspective, l'apport de Duclos et Araar (2006), les applications concernant les courbes de concentration et la taxation indirecte (*Cf*. Makdissi et Mussard, 2006), la pauvreté et les transferts entre les groupes (*Cf*. Makdissi et Mussard, 2006), ou encore le partage de la valeur ajoutée (*Cf*. Mussard et Philippe, 2007).

BIBLIOGRAPHIE

Adams, R. (1994), « Non-Farm Income and Inequality in Rural Pakistan: A Decomposition Analysis », *Journal of Development Studies*, 31(1) : 110-133.

Araar, A. (2002), « L'impact des variations des prix sur les niveaux d'inégalité et de bien-être : une application à la Pologne durant la période de transition », *L'Actualité Économique*, 78(2).

Araar, A. (2006), « On the Decomposition of the Gini Coefficient: An Exact Approach, with an Illustration Using Cameroonian Data », Cahier 06-02 CIRPÉE.

Atkinson, A. (1970), « On the Measurement of Inequality », *Journal of Economic Theory*, 55 : 244-263.

Auvray, C. et A. Trannoy (1992), « Décomposition par source de l'inégalité des revenus à l'aide de la Valeur Shapley », Journées de Microéconomie Appliquée, Sfax, Tunisie.

Cancian, M. et D. Reed (1998), « Assessing the Effects of Wives' Earnings on Family Income Inequality », *Review of Economics and Statistics*, 80(1) : 73-79.

Chantreuil, F. et A. Trannoy (1999), « Inequality Decomposition Values: The Trade-off Between Marginality and Consistency », DP 9924 THEMA.

Chateauneuf, A., T. Gajdos et P.-H. Wilthien (2002), « The Principle of Strong Diminishing Transfer », *Journal of Economic Theory*, 103 : 311-333.

Dagum, C. (1997), « A New Approach to the Decomposition of the Gini Income Inequality Ratio », *Empirical Economics,* 22(4) : 515-531.

Dalton, H. (1920), « The Measurement of Inequality of Incomes », *Economic Journal*, 30 : 348-361.

Datt, G. et M. Ravallion (1992), « Growth and Redistribution Components of Changes in Poverty Measures – A Decomposition with Applications to Brazil and India in the 1980s », *Journal of Development Economics*, 38 : 275-296.

Duclos, J.-Y. et A. Arrar (2006), *Poverty and Equity – Measurement, Policy, and Estimation with DAD*, Springer.

Fei, J.C.H., G. Ranis et S.W.Y Kuo (1978), *Growth with Equity – The Taiwan Case*, Oxford, Oxford University Press.

Foster, J.E., J. Greer et E. Thorbecke (1984), « Notes and Comments. A Class of Decomposable Poverty Measures », *Econometrica*, 52 : 761-766.

Fields, G. (1979), « Income Inequality in Urban Columbia: A Decomposition Analysis », *Review of Income and Wealth*, 25(3) : 327-341.

Flückiger, Y. et J. Silber (1995), « Income Inequality by Income Source and the Breakdown of Inequality Differences Between Two Population Subgroups », *Swiss Journal of Economics and Statistics*, 131(4/1) : 599-615.

Garner, T. (1993), « Consumer Expenditures and Inequality: An Analysis Based on Decomposition of the Gini Coefficient », *Review of Economics and Statistics*, 75(1) : 134-138.

Gastwirth, J.L. (1972), « The Estimation of the Lorenz Curve and the Gini Index », *Review of Economics and Statistics*, LIV : 306-316.

Gini, C. (1912), « Variabilità e mutabilità » *in* Memori di Metodologia Statistica, Vol. 1, Variabilità e Concentrazione. Libreria Eredi Virgilio Veschi, Rome, p. 211-382.

Gini, C. (1914), *L'Ammontare e la composizione della ricchezza delle nazione*, Bocca, Torino.

Gini, C. (1916), « Il concetto di transvariazione e le sue prime applicazioni », dans C. Gini (éd.) (1959), *Giornale degli Economisti e Rivista di Statistica*, p. 21-44.

Keynes, J.M. (1936), *La théorie générale de l'emploi, de l'intérêt et de la monnaie*, traduit de l'anglais par Jean de Largentaye, Payot, Paris.

Kolm, S.-C. (1966), « The Optimal Production of Social Justice », Colloques Internationaux du CNRS, Biarritz, 2-9 septembre 1966.

Kuznets, S. (1955), « Economic Growth and Income Inequality », *American Economic Review*, 45 : 1-28.

Lerman, R. et S. Yitzhaki (1985), « Income Inequalities Effects by Income Source: A New Approach and Applications to United States », *Review of Economics and Statistics*, 67 : 151-156.

Lerman, R. et S. Yitzhaki (1989), « Income Sources and Income Inequality: Measurement from Three US Income Surveys », *Journal of Economic and Social Measurement*, 15(2) : 167-179.

Lorenz, M.O. (1905), « Methods for Measuring Concentration of Wealth », *Journal of the American Statistical Association*, 70 : 209-219.

MAKDISSI, P. et S. MUSSARD (2006), « Decomposition of *s*–Concentration Curves », Cahier 06-41 CIRPÉE, Cahier 06-21 GRÉDI, Working Paper IRISS 2006-09.

MILANOVIC, B. (1997), « A Simple Way to Calculate the Gini Coefficient, and Some Implications », *Economics Letters*, 56 : 45-49.

MORDUCH, J. et T. SICULAR (2002), « Rethinking Inequality Decomposition, with Evidence from Rural China », *Economic Journal*, 112(476) : 93-106.

MUSSARD, S. (2004a), « Décompositions multidimensionnelles du rapport moyen de Gini. Applications aux revenus italiens de 1989 et 2000 », Thèse, Université de Montpellier I.

MUSSARD, S. (2004b), « The Bidimensional Decomposition of the Gini Ratio. A Case Study: Italy », *Applied Economics Letters*, 11(8) : 503-505.

MUSSARD, S. (2006), « Une nouvelle décomposition de la mesure de Gini en sources de revenu et la décomposition en sous-populations : une réconciliation », *Annales d'Économie et de Statistique*, 81 : 1-25.

MUSSARD, S. et B. PHILIPPE (2007), « Une évaluation du rôle des déterminants du partage de la valeur ajoutée », Working Paper IRISS 2007-08.

OWEN, G. (1977), « Values of Games with Priori Unions » *in* R. HEIM et O. MOESCLIN (éds), *Essays in Mathematical Economics and Game Theory*, Springer-Verlag.

PARETO, V. (1896), *Écrits sur la courbe de la répartition de la richesse*, Oeuvres complètes de Vilfredo Pareto publiées sous la direction de Giovanni Busino. Genève, Librairie Droz, 1965.

PYATT, G., C.-N. CHEN et J. FEI (1980), « The Distribution of Income by Factor Component », *Quarterly Journal of Economics*, 95(3) : 451-473.

PODDER, N. (1993), « The Disaggregation of the Gini Coefficient by Factor Components and Its Applications to Australia », *Review of Income and Wealth*, 39(1) : 51-61.

RAO, V.M. (1969), « Two Decompositions of Concentration Ratio », *Journal of the Royal Statistical Society*, Séries A, 132 : 418-425.

SASTRE, M. et A. TRANNOY (2002), « Shapley Inequality Decomposition by Factor Components: Some Methodological Issues », *Journal of Economics*, Supplément 9 : 51-90.

SEN, A.K. (1973), *On Economic Inequality*, Clarendon Press, Oxford.

SHAPLEY, L. (1953), « A Value for N-person Games », *in* H.W. KUHN et A.W. TUCKER (éds), *Contributions to the Theory of Games*, Vol. 2, Princeton University Press.

SHORROCKS, A.F. (1982), « Inequality Decomposition by Factor Component », *Econometrica*, 50 : 193-211.

SHORROCKS, A.F. (1983), « The Impact of Income Component on the Distribution of Family Incomes », *Quarterly Journal of Economics*, 98(2) : 311-326.

SHORROCKS, A.F. (1999), *Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value*, Mimeo, University of Essex.

SILBER, J. (1989), « Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality », *Review of Economics and Statistics*, 71 : 107-115.

SILBER, J. (1993), « Inequality Decomposition by Income Source: A Note », *Review of Economics and Statistics*, 75(3) : 545-547.

SILBER, J. (2004), « Inequalities: Theory, Experiments and Applications », *European Journal of Political Economy*, 20 : 813-820.

THEIL, H. (1967), *Economics and Information Theory*, North-Holland Publishing Company, Amsterdam.

TSUI, K. (1999), « Multidimensional Inequality and Multidimensional Generalized Entropy Measures: An Axiomatic Derivation », *Social Choice and Welfare*, 16 : 145-157.

WDR (2006), *Rapport sur le développement dans le monde 2006 – Équité et Développement*, Banque Mondiale.

YITZHAKI, S. (1983), « On an Extension of the Gini Inequality Index », *International Economic Review*, 24(3) : 617-628.

YITZHAKI, S. (1990), « On the Effect of Subsidies to Basic Commodities on Inequality in Egypt », *Oxford Economic Papers*, 42 : 772-792.